Wang, L., Parnell, A., Williams, C., Bakar, N. A., Challand, M. R., van der Kamp, M. W., Simpson, T. J., Race, P. R., Crump, M. P., & Willis, C. L. (2018). A Rieske oxygenase/epoxide hydrolase-catalysed reaction cascade creates oxygen heterocycles in mupirocin biosynthesis. *Nature Catalysis*, *1*(12), 968-976. https://doi.org/10.1038/s41929-018-0183-5

# A Rieske oxygenase / epoxide hydrolase-catalysed reaction cascade creates oxygen heterocycles in mupirocin biosynthesis

Luoyi Wang,[1,3] Alice Parnell,[2,3] Christopher Williams,[1,3] Nurfarhanim A. Bakar,[1] Martin R. Challand,[2,3] Marc W. van der Kamp,[2,3] Thomas J. Simpson,[1,3] Paul R. Race,[2,3] Matthew P. Crump,[1,3*] Christine L. Willis [1,3*]

[1] School of Chemistry, Cantock's Close, University of Bristol, BS8 1TS, UK

[2] School of Biochemistry, University Walk, University of Bristol, BS8 1TD, UK

[3] BrisSynBio Centre for Synthetic Biology Research, Tyndall Avenue, University of Bristol, BS8 1TQ, UK

[*]E-mail: matt.crump@bristol.ac.uk; chris.willis@bristol.ac.uk

## Abstract

Oxygen heterocycles, in particular tetrahydropyrans and tetrahydrofurans, are common structural features of many biologically active polyketide natural products. Mupirocin is a clinically important antibiotic isolated from *Pseudomonas fluorescens* and is assembled on a tetrahydropyran ring which is essential for bioactivity. However, the biosynthesis of this moiety has remained elusive. Here we show an oxidative enzyme-catalysed cascade that generates the tetrahydropyran ring of mupirocin. A Rieske non-heme oxygenase (MupW) catalysed selective oxidation of the C8-C16 single bond in a complex acyclic precursor is combined with an epoxide hydrolase (MupZ) to catalyse the subsequent regioselective ring formation to give the hydroxylated tetrahydropyran. In the absence of MupZ, a 5-membered tetrahydrofuran ring is isolated and model studies are consistent with cyclisation occurring via an epoxide intermediate. High resolution X-ray crystallographic studies, molecular modelling and mutagenesis experiments of MupZ provide insights into tetrahydropyran ring formation proceeding via an anti-Baldwin 6-*endo*-*tet* cyclisation.

Over many decades if not centuries, drug discovery has been inspired by natural products derived from microbes, plants and animals[1]. Natural products not only display valuable biological properties but often have complex and architecturally unique scaffolds. Elucidation of natural product biosynthetic pathways provides important insights into the assembly of potent bioactive molecules and often reveals unprecedented chemical and enzymatic platforms enabling the genetic manipulation of pathways to provide new compounds with improved properties towards stability and/or bioactivity[2,3].

Mupirocin, a mixture of polyketide-derived congeners consists of pseudomonic acids A (PA-A, **1**, Fig. 1a), B (**2**), and C (**3**) isolated from *Pseudomonas fluorescens* NCIMB 10586, is a clinically important antibiotic active against Gram-positive bacteria, including methicillin resistant *Staphylococcus aureus* (MRSA)[4,5]. It remains one of the most-effective treatments for MRSA, used mainly for topical skin infections and as a pre-operative nasal spray to combat hospital borne *S. aureus* infections. The 6,7-dihydroxy-tetrahydropyran (THP) ring is essential for its antibiotic activity. Classical labelling studies and more recent genetic studies have established the biosynthetic origins of mupirocin[6-10]. It contains a polyketide moiety produced by a *trans*-AT modular polyketide synthase esterified by an unusual 9-hydroxynonanoic acid moiety. The length of the fatty acid chain has a strong effect on its activity, which is due to inhibition of the isoleucyl tRNA synthetase of inter alia, *E. coli* and *S. aureus*[11]. Oxygen-18 labelling has shown that the essential THP ring is formed by ether linkage from the 5-OH to the 16-methyl group in a linear precursor (see e.g. Fig. 1c)[6]. A priori there are a number of chemically and biosynthetically precedented mechanisms for this cyclisation: electrophilic addition to an 8,16-alkene, conjugate addition onto an 8,16-dien-7-one, direct nucleophilic displacement on an activated derivative of the 16-methyl, or addition of the 5-OH to an 8,16-epoxide. The latter mechanism is given credence by the recent demonstration that PA-B, with an extra 8-hydroxyl group, is actually an intermediate in PA-A biosynthesis, being converted to PA-A in a complex sequence controlled by 7 genes, involving 7 discrete intermediates, two of which are carrier protein-bound[8-10]. Oxygen heterocycles are widespread structural features of many biologically active polyketide natural products and gaining an understanding of how tetrahydropyrans (THPs) and tetrahydrofurans (THFs) are formed is of wide current interest world-wide[12-14].

We have previously shown that deletion of the dioxygenase MupW (and also of its redox partner MupT) in the mupirocin gene cluster abolished production of the PAs and led to accumulation of mupirocin W1 (**6**) which lacks the THP ring, but which now contains a new THF ring (Fig. 1b). It is proposed that the true biosynthetic intermediate is the linear compound **5** which undergoes spontaneous intra-molecular attack by the 7-OH onto the 10,11-epoxide[9,15] (Fig. 1b). In support of this hypothesis, double inactivation of *mupW* and the oxidoreductase (OR) encoded by *mmpE* (identified

as responsible for 10,11-epoxidation) results in accumulation in the resulting *mmpEΔOR/ΔmupW* mutant of *P. fluorescens*, albeit in very low titres of the much more stable, linear

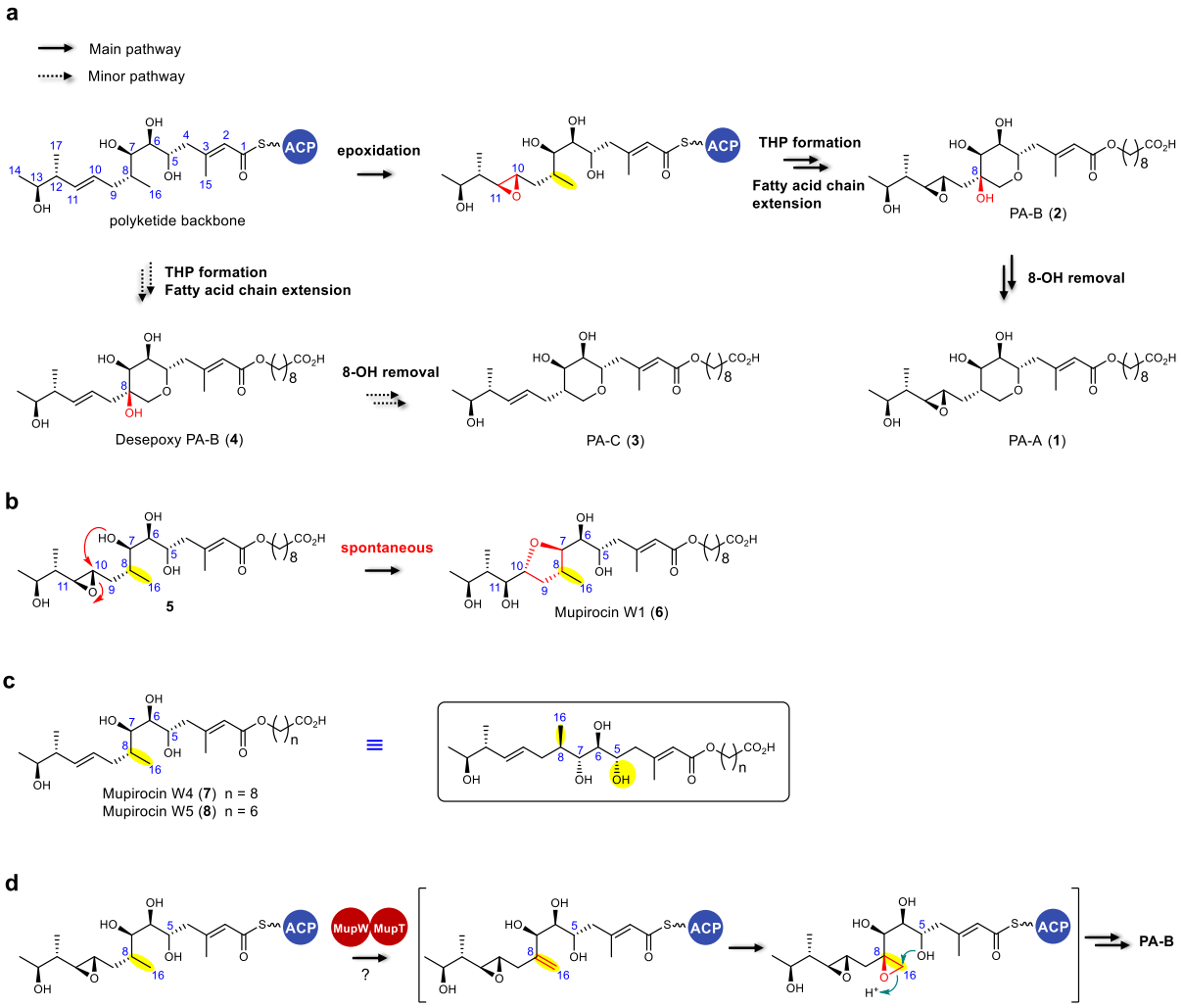


**Figure 1. Summary of mupirocin biosynthesis in *P. fluorescens* NCIMB 10586**. **a**, Key post-assembly steps in mupirocin biosynthesis. **b**, Structures of the proposed biosynthetic intermediate **5** which undergoes spontaneous cyclisation to mupirocin W1 (**6**) isolated from *ΔmupW* mutant. **c**, Structures of mupirocin W4 (**7**) and W5 (**8**) isolated from a *mmpEΔOR/ΔmupW* double mutant. **d**, Previously proposed mechanism of THP formation in mupirocin biosynthesis[9]. THP, tetrahydropyran.

metabolites, mupirocin W4 (**7**) and W5 (**8**)[9] (Fig. 1c). Thus, MupW and MupT appeared to be key enzymes in THP formation and we proposed that the oxygenase MupW together with its associated ferredoxin MupT catalyse oxidation of the C8-C16 single bond to an epoxide, which would then be opened by the 5-OH to form the THP ring and simultaneously install the 8-OH[9] (Fig 1d). However, experimental proof for this is lacking and whether subsequent cyclisation to give the THP is a spontaneous process or if this reaction is also mediated by another enzymatic activity remained unclear.

Intrigued by nature's ability to selectively oxidise the non-activated C8-C16 single bond despite the presence of functional groups (double bonds and hydroxyl groups) elsewhere in the molecule (Fig. 1c,

boxed structure), herein we describe the reactions catalysed by MupW on mupirocin W4 (**7**) and W5 (**8**) and show that formation of the THP ring requires the dual action of the dioxygenase MupW and an epoxide hydrolase MupZ, which had not previously been assigned a function. In the absence of MupZ, MupW catalyses formation of a 5-membered tetrahydrofuran and model studies are consistent with cyclisation occurring via an epoxide intermediate. High resolution X-ray crystallographic studies, molecular modelling and mutagenesis experiments of MupZ provide insights into tetrahydropyran ring formation proceeding via an anti-Baldwin *6-endo-tet* cyclisation.

## Results

### MupW is a Rieske non-heme oxygenase

Bioinformatic analysis suggested that MupW is a Rieske oxygenase-like enzyme in which the *N*-terminal domains contain a [2Fe-2S] Rieske cluster involved in electron transfer reactions from a ferredoxin shuttle protein to a non-heme iron centre in the C-terminal domain. The conserved sequence motifs CXH and CXXH present two Cys and two His residues that ligate the iron atoms of the [2Fe-2S] cluster[16]. Within the *C*-terminal domain, a conserved $DX_2HX_4H$ motif contains two His residues that ligate a non-heme iron centre and an aspartate residue[16]. All three universally conserved sequence motifs are found within the MupW protein with the exception that aspartate is substituted for glutamate but which has been shown to be functionally equivalent in other Rieske non-heme like oxygenases[17] (Fig. 2a).

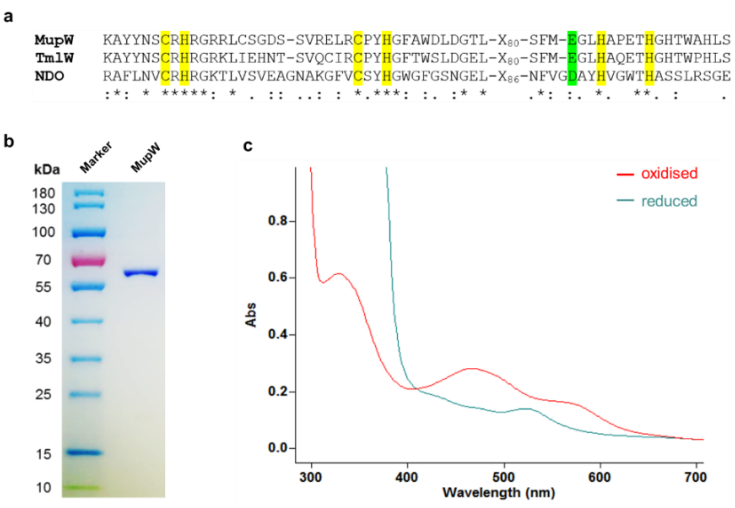


**Figure 2. Characterisation of the Rieske non-heme oxygenase MupW. a**, Sequence alignment of the Rieske non-heme oxygenases MupW and TmlW (the MupW equivalent in thiomarinol pathway) with the structurally characterised Rieske oxygenase naphthalene dioxygenase (NDO), conserved residues that ligate the [2Fe-2S] cluster and Fe atom are highlighted in yellow. **b**, SDS-PAGE of $His_6$-tagged recombinant MupW. **c**, UV absorption spectrum of purified MupW (4 mg/mL) in oxidised form (red line) and sodium dithionite-reduced form (blue line).

The *N*-terminal $His_6$-tagged recombinant MupW was heterologously overexpressed and purified to homogeneity (Fig. 2b) from *Escherichia coli* BL21 (DE3) to yield a brown-red enzyme solution that was characterised by ultraviolet-visible (UV) spectroscopy using standard techniques. The aerobically purified enzyme solution displayed characteristic [2Fe-2S] cluster absorbance peaks at around 330,

460, and 560 nm, while upon reduction with sodium dithionite anaerobically, new absorption maxima were observed at around 520 nm[18] (Fig. 2c).

Rieske non-heme iron oxygenases are typically multicomponent enzymes that utilise a reductase and a ferredoxin partner for electron transfer from NAD(P)H to the terminal oxygenase that performs catalysis[19]. Sequence alignment and the UV spectrum (Supplementary Figure 1) suggests that MupT is the [2Fe-2S] Rieske cluster containing ferredoxin component, however, no cognate reductase partner has been identified within the mupirocin gene cluster[7,8]. Due to this limitation, $O_2$ saturated putative substrates mupirocin W4 (**7**) and W5 (**8**) were then incubated with sodium dithionite-reduced MupW anaerobically, however, no turnover was observed by liquid chromatography–mass spectrometry (LC-MS) analysis potentially due to extremely low yield of products (data not shown).

## Whole-cell biotransformations reveal MupW catalysed oxidations

We next turned to in vivo biotransformations. Substrates were incubated with whole live cells of the MupW overexpressing *E. coli* strain to obviate the redox partners and cofactor dependency. Such an approach has been widely used[20-23] and was elegantly exemplified by Hauer and co-workers in their work on engineering Rieske non-heme iron oxygenases for the asymmetric dihydroxylation of alkenes[21]. Following Hauer's protocol, we examined reactions of mupirocin W4 (**7**) and W5 (**8**) in whole-cell biotransformations using resting MupW overexpressing *E. coli* cells supplemented with glucose for in situ NAD(P)H regeneration. After incubation, mupirocin W4 (**7**, *m/z* 485 $[M-H]^-$, Fig. 3a including appropriate controls) was fully converted to a product, identical in mass (*m/z* 499 $[M-H]^-$) and HPLC retention time (11.9 mins) with those expected for desepoxy PA-B (**4**), apparently consistent with THP formation. When incubated with cells containing only the empty pET28a vector or denatured cells, the substrate mupirocin W4 (**7**) remains intact and no new product was detected (Fig. 3a iv), confirming that the MupW protein is indeed responsible for formation of the product.

1D and 2D NMR spectra of the new product were, however, found to be significantly different from those of desepoxy PA-B (**4**) (Fig. 3b, Supplementary Figure 2). The structure of the new compound (**9**) was established by HRESI-MS and detailed analysis of 2D NMR data, including $^1$H-$^1$H COSY, HSQC, HMBC and NOESY as shown in Fig. 3b and was determined to have a five membered-tetrahydrofuran (THF) ring formed between 5-OH and C8, and a hydroxymethyl group (C16) attached to C8 (Supplementary Figure 3, Supplementary Table 1). Mupirocin W5 (**8**), with a shorter fatty acid chain than **7**, was also fully converted to compound **10** assembled on a tetrasubstituted THF (Supplementary Figure 4). The structure of **10** was further supported by chemical derivatisation to afford acetonide **11**, which was further acetylated to yield **12**, confirming the presence of the 8-hydroxymethyl group (Fig. 3c, Supplementary Methods).

To gain further insight into formation of the unexpected tetrahydrofurans a model study was undertaken. Treatment of alkene **13** with *meta*-chloroperbenzoic acid (*m*CPBA) was predicted to give a mixture of epoxides but these were not detected as spontaneous cyclisation occurred giving a 7:3 mixture of diastereomeric tetrahydrofurans **14** and **15** in 80% yield (Fig. 3d, Supplementary Methods). Hence, formation

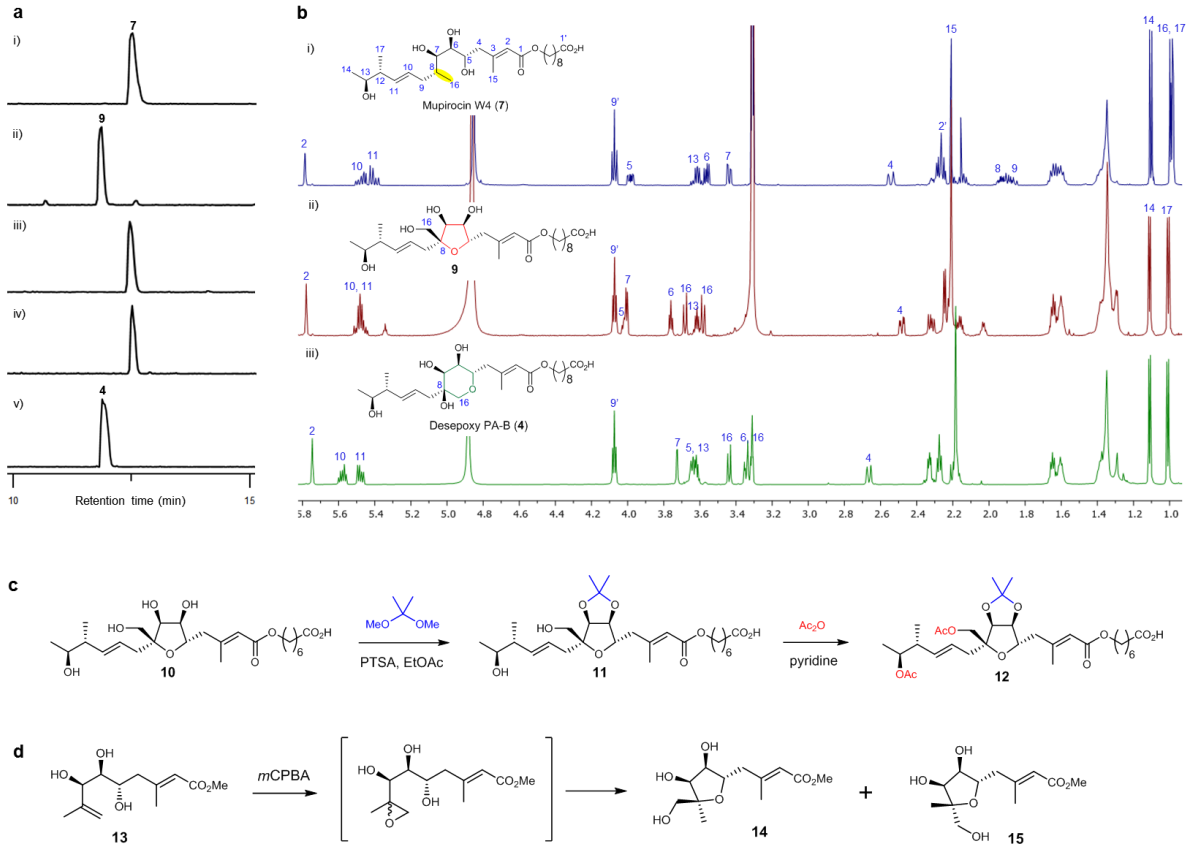


**Figure 3. Whole-cell biotransformations reveal MupW catalysed oxidations**. **a**, LC-MS (total ion current, negative mode) traces of (i) mupirocin W4 (**7**), (ii) mupirocin W4 (**7**) incubated with MupW producing whole cells at 30 °C for 6 hours, (iii) mupirocin W4 (**7**) incubated with denatured MupW cells (boiled at 95 °C for 10 min), (iv) mupirocin W4 (**7**) incubated with cells harbouring the empty pET28a vector, (v) desepoxy PA-B (**4**) standard. **b**, $^{1}$H NMR spectra of (i) mupirocin W4 (**7**), (ii) new product **9**, (iii) desepoxy PA-B (**4**) standard. **c**, Chemical derivatisation of compound **10. d**, Model reaction showing spontaneous cyclisation to 5-membered ring products.

of tetrahydrofurans **9** and **10** from incubation of **7** and **8** with MupW producing whole cells, is consistent with spontaneous cyclisation of an 8,16-epoxide. Epoxide ring opening to give Baldwin or anti-Baldwin behaviour has been the subject of detailed investigations[24-26]. In general, ring opening of $\alpha,\beta$-epoxides by a $\delta$-hydroxyl is predicted to give a THF ring rather than a 6-membered THP. Indeed, as discussed above, we have previously found in studies on mupirocin biosynthesis spontaneous cyclisation of epoxide **5** occurs to give tetrahydrofuran **6**[15] (Fig. 1b).

To investigate if tetrahydrofuran **9** is an intermediate in mupirocin biosynthesis, it was fed to cultures of the *mmpEΔOR/ΔmupW* mutant of *P. fluorescens* but returned unchanged starting material. In

contrast, tetrahydropyran desepoxy PA-B (**4**) was readily transformed to PA-C (**3**) in the same mutant (Supplementary Figure 5). Thus, it is highly likely that the THF ring products **9** and **10** are artefacts spontaneously formed from the corresponding epoxides and are not intermediates in mupirocin biosynthesis. Hence THP ring formation must be an enzyme-controlled process.

## MupZ catalyses formation of the THP ring

We thus turned our attention to identifying the enzyme responsible for the regioselective attack of the 5-hydroxy to C-16 to generate the 6-membered THP ring. We have previously identified the functions of most of the genes in the mupirocin gene cluster[7,8] but several open reading frames with putative or unassigned functions remain. Prominent among the latter group was MupZ (Fig. 4a). Bioinformatic analysis indicated MupZ as a mixed α/β protein[27], but its role in mupirocin biosynthesis had not been investigated.

To investigate whether MupZ is indeed the missing epoxide hydrolase, whole-cell biotransformations were performed by incubating *E. coli* co-transformed with *mupZ* and *mupW* (Fig. 4b) with linear precursors. Incubation with mupirocin W4 (**7**) again yielded a product with the same retention time (11.9 mins) and mass (*m/z* 499 $[M-H]^-$) as desepoxy PA-B (**4**) (Fig. 4c) and in this instance NMR spectroscopy confirmed this was indeed desepoxy PA-B (**4**) (Supplementary Figure 6). Similarly, mupirocin W5 (**8**) was fully converted to compound **16** containing a THP ring as opposed to the formation of product **10** with a THF ring catalysed by MupW alone (Supplementary Figure 6). Control reactions with denatured cells or with cells only expressing MupZ resulted in no product formation. To further confirm the function of MupZ, gene disruption of this enzyme within the *mmpEΔOR* mutant of *P. fluorescens* NCIMB 10586 was carried out using the two-step allelic exchange method (Supplementary Figure 7)[28]. The *mmpEΔOR* mutant of *P. fluorescens* produces mainly PA-C (**3**) with a 10,11-double bond, rather than the less stable PA-A (**1**) with the 10,11-epoxide, albeit in low yield. Deletion of *mupZ* from the *mmpEΔOR* mutant abolished PA-C production and resulted in production of the two THF ring compounds **9** and **10**, which were named mupirocin Z1 and Z2, respectively, and trace amounts of the acyclic metabolites mupirocin W4 (**7**) and W5 (**8**) (Fig. 4d). *In trans* expression of the *mupZ* gene in this mutant restored PA-C production (Fig. 4d).

These results demonstrate that MupW catalyses the oxidation of the C8-C16 bond of acyclic mupirocins W4 (**7**) and W5 (**8**) to substrates for MupZ which subsequently catalyses regioselective cyclisation to generate the THP ring in mupirocin biosynthesis. Model studies are in accord with the oxidised intermediates being epoxides (Fig 3d) and our next goal was to gain further evidence by determining the structure of MupZ using X-ray crystallography and then to use molecular modelling and docking studies with the proposed substrate.

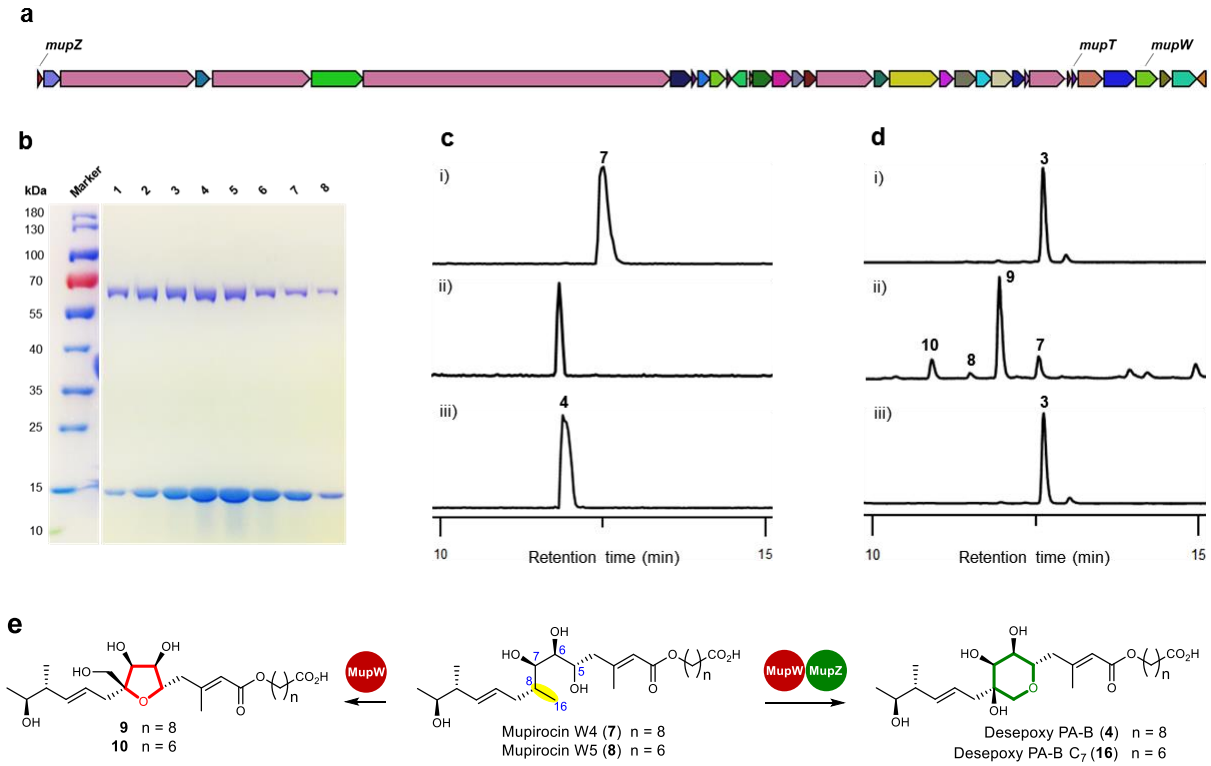


**Figure 4. MupZ catalyses formation of the THP ring**. **a**, Organisation of mupirocin gene cluster, see ref 5 for full annotation. **b**, SDS-PAGE analysis of co-expression of MupW and MupZ, lanes 1-8: fractions from nickel column purification. **c**, LC-MS (total ion current, negative mode) traces of (i) mupirocin W4 (**7**), (ii) product formed by incubating mupirocin W4 (**7**) with *E. coli* cells overexpressing both MupW and MupZ, and (iii) desepoxy PA-B (**4**). **d**, HPLC traces of (i) *mmpEΔOR* mutant of *P. fluorescens* NCIMB 10586, (ii) *mmpEΔOR/ΔmupZ* double mutant, and (iii) *mmpEΔOR/ΔmupZ* complemented with mupZ *in trans*. **e**, Structures of products isolated from biotransformations with MupW and MupW/MupZ.

The structure of MupZ was determined to 1.45 Å resolution by molecular replacement using the conserved protein MM_1583 from *Methanosarcina mazei* Go1 (PDB ID: 4DPO, 16% identity) as a search model and refined to an $R_{work}/R_{free}$ of 16.1/19.5 (PDB ID: 6FXD, Supplementary Table 2). The asymmetric unit comprises two molecules of MupZ arranged as a symmetric homodimer (Fig. 5a). The enzyme adopts an α+β barrel-fold, formed from a central 5 stranded anti-parallel β-sheet augmented by three α-helices. DALI analysis[29] of the MupZ structure reveals the absence of any closely structurally related (Z score >2.0) epoxide hydrolases in the Protein Data Bank (PDB). The only related enzymes identified were Lsd19 (PDB ID: 3RGA, $C_\alpha$ RMSD = 4.9), from the lasalocid A biosynthetic pathway[30], and MonBI (PDB ID: 3WMD, $C_\alpha$ RMSD = 9.0), from the monensin pathway[31]. Both proteins are α+β barrel enzymes involved in epoxide ring opening cascades during the biosynthesis of polyether natural products. Neither exhibits significant structural identity to MupZ beyond retention of a core barrel fold.

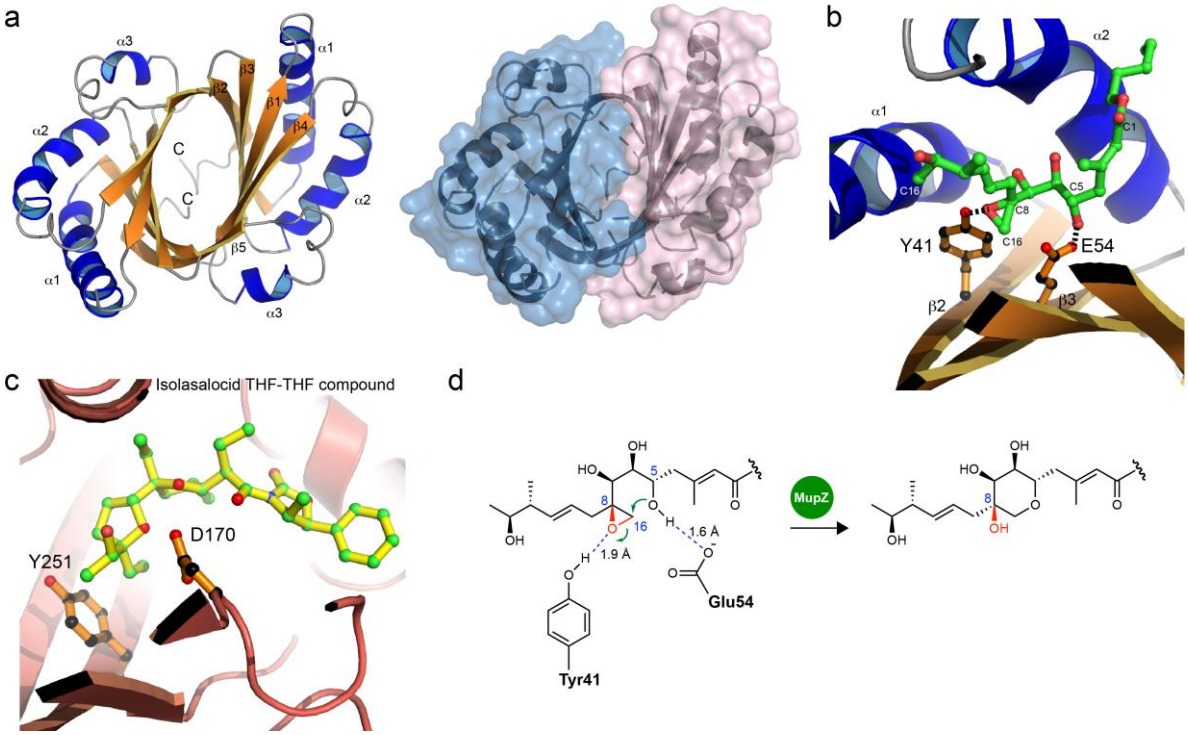


**Figure 5. The X-ray structure of MupZ (PDB ID: 6FXD) and proposed mechanism of cyclisation in mupirocin biosynthesis. a**, Overall fold of MupZ dimer. **b**, Docked structure of putative epoxide substrate in the active site of MupZ. **c**, Structure of Lsd19 (PDB ID: 3RGA) with the isolasalocid-like THF-THF ligand in the active site[30]. **d**, Proposed mechanism of epoxide-opening cyclisation/THP formation in mupirocin biosynthesis.

The epoxide hydrolase Lsd19B has been reported to utilise an Asp–Tyr dyad as a general acid–base pair to catalyse epoxide-opening to form cyclic ethers[30]. Docking and molecular modelling studies of MupZ with the proposed epoxide substrate identified an analogous acid–base pair of Glu54-Tyr41 in the enzyme active site (Fig. 5b and 5c). To investigate the role of these two residues the point mutants MupZ-Y41F and MupZ-E54Q were produced. Co-expression of MupW together with either MupZ-Y41F or MupZ-E54Q was successfully achieved using the same protocol as for the wild type MupZ (Supplementary Figure 8). Whole-cell biotransformation of substrates mupirocin W4 (**7**) or W5 (**8**) by MupZ-Y41F or MupZ-E54Q led to the formation of the 5-membered THF ring products in contrast to the THP ring products produced by WT-MupZ (Supplementary Figure 8). Together these data demonstrate the importance of these residues for catalysis. On the basis of these findings, we propose an acid–base mechanism of epoxide ring opening by MupZ where Glu54 acts to deprotonate the C5 hydroxyl group of the substrate followed by protonation of the epoxide oxygen by Tyr41 and subsequent stabilisation of the developing transition leading to the 8-hydroxy THP ring structure (Fig. 5d).

## Discussion

We have identified a tandem Rieske non-heme oxygenase/epoxide hydrolase (MupW/MupZ) that catalyses THP formation in the biosynthesis of the clinically important antibiotic mupirocin. These intriguing transformations couple the MupW catalysed selective oxidation of the non-activated C8-C16 bond in complex acyclic precursors, mupirocins W4 (**7**) and W5 (**8**) to deliver substrates for MupZ which, in turn, catalyses regioselective cyclisation between the 5-OH and C-16 to generate the THP ring of desepoxy PA-B (**4**) and the analogue with a shorter fatty acid side-chain (**16**). In the absence of MupZ, the acyclic substrates **7** and **8** were transformed to tetrahydrofurans **9** and **10**, with cyclisation occurring between 5-OH and C-8. Model chemical reactions as well as molecular modelling/docking studies are in accord with the acyclic biosynthetic precursors **7** and **8** being transformed to 8,16-epoxides prior to ring formation via an anti-Baldwin *6-endo*-tet ring closure catalysed by MupZ to generate the 8-hydroxy tetrahydropyran ring.

There are two features of this enzyme-catalysed sequence which are particularly intriguing. First, chemically challenging functionalisation of non-activated bonds in the presence of numerous proximal functional groups is occurring with exquisite selectivity catalysed by the Rieske non-heme oxygenase, MupW. These enzymes are more usually associated with dihydroxylation of arenes and alkenes[21], oxidative transformations of aliphatic compounds including monohydroxylations[32] and desaturation[33] (probably via dehydration of a mono-hydroxylated species) of non-activated C-H groups and oxidative ring closures via a putative radical activation of an aliphatic C-H centre[17]. Further studies will be required to definitively prove that an epoxide is generated by MupW, which is a very unusual overall transformation for a Rieske non-heme oxygenase catalysed process. In contrast, α-ketoglutarate-dependent (α-KG) oxygenases are known to give epoxides e.g. H6H is an α-KG oxygenase that hydroxylates (−)-atropine and then closes the newly introduced oxygen onto a neighbouring methylene to generate the epoxide of scopolamine[34].

The second unusual feature of tetrahydropyran formation in mupirocin biosynthesis is the anti-Baldwin 6-*endo*-tet cyclisation catalysed by the epoxide hydrolase MupZ giving the hydroxy tetrahydropyran, desepoxy PA-B (**4**). There are very few examples where the nominally disfavoured THP has been shown to be the preferred product of an epoxide hydrolase enzyme. The flavin-dependent monooxygenase (FMO) and an epoxide hydrolase pair AurC/AurD iteratively transform a terminal triene into the dioxabicyclo-octane system in the aurovertin fungal metabolite (Fig. 6a)[35]. Furthermore, the biosynthesis of the polyether lasalocid involves the FMO/epoxide hydrolase pair Lsd18/Lsd19 in which a 6-*endo*-tet cyclisation of epoxy-alcohol **17** generates the hydroxylated tetrahydropyran[30,36] (Fig. 6a). High-resolution X-ray crystal structure of Lsd19 (also known as LasB) together with quantum-mechanical calculations studies have suggested a mechanism of general acid–base catalysis and shown that *5-exo*-tet cyclisation is energetically favoured under both acid- and base

catalysis[30]. Indeed, competing 5-*exo*-tet cyclisations have proved problematic in the preparation of substrates for Lsd19 in lasalocid biosynthesis[37] as well as in our work on the synthesis of epoxides as putative substrates for MupZ.

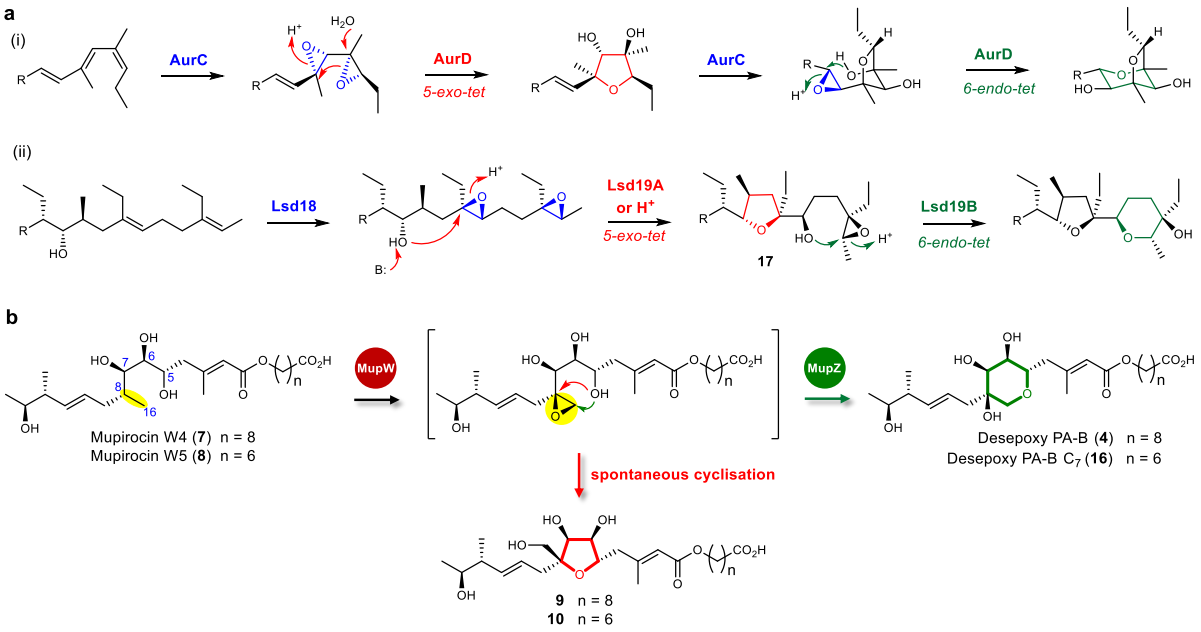


**Figure 6. Enzyme-catalysed reactions giving hydroxylated tetrahydropyrans via 6-*endo*-tet cyclisations. a,** Examples of FMO/epoxide hydrolase cascades in (i) aurovertin[35] and (ii) lasalocid[30,36] biosynthesis. **b,** Formation of oxygen heterocycles in mupirocin biosynthesis**.**

Based on the X-ray crystal structure of MupZ, molecular modelling and mutagenesis experiments, we propose that the acid-base pair Glu54 and Tyr41 are required in the active site to catalyse formation of the 8-hydroxytetrahydropyran ring in mupirocin biosynthesis (Fig. 5d). Tetrahydrofuran formation in our model studies combined with the functional characterisation of the catalytic dyad in MupZ provide support for an epoxide intermediate (Fig. 6b) generated by MupW from a non-activated alkane, an unusual overall transformation not previously reported for Rieske oxygenases[38].

## Methods

**General experimental procedures.** Column chromatographic separations were carried out by using MCI gel CHP20P (75-150 µm, Mitsubishi Chemical Corporation, Tokyo, Japan) or Sephadex LH-20 (Pharmacia Biotech AB, Uppsala, Sweden) as packing materials. LC-MS data were obtained on a Waters LCMS system comprising Waters 2767 autosampler, Waters 515 HPLC pump, Waters 2998 Diode Array detector, Waters 2424 ELS detector and Waters Quatro Micro mass spectrometer. HPLC grade $H_2O$ and MeCN were added with 0.05% formic acid as solvent system. Analytical LC-MS data were obtained using a Phenomenex Kinetex column ($C_{18}$, 250 x 4.60 mm, 5 µm) at a flow rate of 1 mL/min. Preparative HPLC purification were carried out using a Phenomenex Kinetex column ($C_{18}$, 250 x 21.20 mm, 5 µm) at a flow rate of 16 mL/min. HRESIMS data were obtained on a Bruker Daltonics micrOTOF II instrument. $^1$H and 2D NMR data were collected on Bruker Cryo500 or Bruker Cryo700 Micro-Coil NMR spectrometers, $^{13}$C NMR data were collected on Bruker Cryo500 NMR spectrometer. For NMR spectra of the compounds in this article, see Supplementary Figures 9-48.

**Isolation and purification of mupirocin W4 (7) and W5 (8) from *mmpEΔOR/ΔmupW* double mutant of *P. fluorescens* NCIMB 10586.** A 1.0 L scale fermentation of *mmpEΔOR/ΔmupW P. fluorescens* NCIMB 10586 was carried out as per general procedure that has been described previously[10]. The crude extract was subjected to MCI gel CHP20P and then Sephadex LH-20 column chromatography and eluted with MeOH to give 5 fractions. Fraction 5 was further purified by HPLC eluted with a gradient of 20 to 70% of MeCN in water over 20 minutes to yield mupirocin W4 (**7**, 1.0 mg) and mupirocin W5 (**8**, 1.5 mg). Mupirocin W4 (**7**): ESI-MS: *m/z* = 485 $[M-H]^-$; $^1$H NMR (500 MHz, $CD_3OD$) and $^{13}$C NMR (125 MHz, $CD_3OD$) data are shown in Supplementary Table 1. Mupirocin W5 (**8**): ESI-MS: *m/z* = 457 $[M-H]^-$; $^1$H NMR (500 MHz, $CD_3OD$): $\delta_H$ 5.79 (1H, br s, H-2), 5.47 (1H, m, H-10), 5.40 (1H, dd, *J* = 15.4, 7.7 Hz, H-11), 4.08 (2H, t, *J* = 6.6 Hz, H-7'), 3.98 (1H, m, H-5), 3.62 (1H, dq, *J* = 6.3, 4.9 Hz, H-13), 3.56 (1H, dd, *J* = 8.6, 4.9 Hz, H-6), 3.44 (1H, dd, *J* = 8.6, 2.9 Hz, H-7), 2.54 (1H, d, *J* = 14.3 Hz, H-4), 2.31 (1H, m, H-9), 2.29 (1H, m, H-4), 2.26 (2H, m, H-2'), 2.21 (3H, s, H-15), 2.15 (1H, m, H-12), 1.93 (1H, m, H-8), 1.88 (1H, m, H-9), 1.66 (2H, m, H-6'), 1.62 (2H, m, H-3'), 1.35-1.45 (4H, m, H-4' and H-5'), 1.10 (3H, d, *J* = 6.4 Hz, H-14), 0.99 (3H, d, *J* = 6.9 Hz, H-17), 0.98 (3H, d, *J* = 6.7 Hz, H-16); $^{13}$C NMR (125 MHz, $CD_3OD$): $\delta_C$ 179.0 (C, C-1'), 168.4 (C, C-1), 159.6 (C, C-3), 134.5 (CH, C-11), 131.3 (CH, C-10), 118.5 (CH, C-2), 78.3 (CH, C-7), 75.1 (CH, C-6), 72.5 (CH, C-5), 72.3 (CH, C-13), 64.7 ($CH_2$, C-7'), 45.3 (CH, C-12), 44.1 ($CH_2$, C-4), 36.2 (CH, C-8), 35.4 ($CH_2$, C-2'), 34.3 ($CH_2$, C-9), 29.9 ($CH_2$, C-5'), 29.7 ($CH_2$, C-6'), 26.9 ($CH_2$, C-4'), 26.2 ($CH_2$, C-3'), 20.2 ($CH_3$, C-14), 19.2 ($CH_3$, C-15), 17.2 ($CH_3$, C-16), 16.6 ($CH_3$, C-17).

**Expression and purification of proteins.** The *mupW* and *mupZ* genes were amplified from *P. fluorescens* NCIMB 10586 genomic DNA by polymerase chain reaction (PCR), with the following primers (lower case letters indicate appropriate sequences homologous to target vectors): *mupW*-For: 5'-tgggtcgcggatccgaattcATGAAAAGCAAATTGATCGAGC-3'; *mupW*-Rev: 5'-caagcttgtcgacggagctcT-CAATCGCGCCTGTCAAGATA-3'; *mupZ*-For: 5'-aagttctgtttcaggcccgATGAATCGCACCTGCATGGCCATGCC-3'; *mupZ*-Rev: 5'-atggtctagaaagctttaCGCCACCACAGACGGCCGGCTG-3'. The *mupW* and *mupZ* PCR products were then introduced into pET28a vector (pre-linearized with EcoRI and SacI), and pOPINF (pre-linearized with KpnI and HindIII), respectively, using the In-Fusion HD Cloning Kit (Clonetech). The resultant plasmids were verified by sequencing and transformed into *E. coli* BL21 (DE3). For MupW, the resultant strain was inoculated into 20 ml LB medium supplemented with 100 µg/ml kanamycin and incubated overnight at 37 °C/200 rpm. The overnight culture was added into 1 L of auto induction medium LB broth base including trace elements (FORMEDIUM™) containing 100 µg/ml kanamycin and incubated at 37 °C/200 rpm until the optical density of the cultures at 600 nm had reached 0.6, and following incubation was changed to 18 °C/200 rpm for overnight. For MupZ, LB broth medium supplemented with 100 µg/ml carbenicillin was used instead of auto induction medium, protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM until the optical density of the cultures at 600 nm had reached 0.6, followed by incubation at 18 °C/200 rpm for overnight. Co-expression of MupW and MupZ followed the same procedure as MupW but with auto induction medium supplemented with both kanamycin and carbenicillin. Synthetic DNAs were purchased from Thermo Fisher Scientific for MupZ-Y41F and MupZ-E54Q point mutants, cloning and

expression followed the same protocol as for the wild type. Cells were then harvested by centrifugation (4 °C, 4500 rpm for 30 min) and resuspended in Buffer A (50 mM Tris/HCl pH 7.5, 150 mM NaCl). The cells were lysed by sonication (6 × 10 s pulsed cycle) on ice, and the debris was removed by centrifugation (4 °C, 15000 rpm for 30 min). The clarified supernatant was loaded onto a pre-equilibrated Ni-NTA agarose column, which was then washed with a gradient of Buffer B (50 mM Tris/HCl pH 7.5, 150 mM NaCl, 1 M imidazole). Fractions containing the target protein of interest, as established by monitoring the absorbance of the column eluent at 280 nm followed by SDS-PAGE analysis, were pooled and concentrated. Concentrated protein samples were subjected to Superdex 75 or 200 size exclusion column for further purification, eluted with Buffer A to yield final protein samples with >95% purity.

**Whole-cell biotransformation**. In *in vivo* biotransformations, 50 mL of overnight culture of *E. coli* BL21 (DE3) cells overexpressing MupW or both MupW and MupZ in auto induction medium were centrifuged at 6000 rpm for 5 min, cell pellets were resuspended in 2 mL of 100 mM potassium phosphate buffer pH 7.2 supplemented with 20 mM glucose, 0.4 mg of substrates mupirocin W4 (**7**) or W5 (**8**) dissolved in 40 µl of MeOH were then added and the reactions were incubated at 30 °C, 180 rpm for 6 h. As negative controls, denatured corresponding cells or cells harbouring empty vectors were used. Reactions were quenched by adding equal volume of acetonitrile, vortexed and centrifuged. The acetonitrile layer was injected for LC-MS analysis. Full conversions from substrates to products were observed according to HPLC analysis. Products were cleaned by preparative HPLC prior to NMR analysis. Compound **9** (0.2 mg): HR-ESI-MS: *m/z* = 523.2861 $[M+Na]^+$ (523.2878 calcd for $C_{26}H_{44}O_9Na$)**;** $^1H$ NMR (700 MHz, $CD_3OD$) and $^{13}C$ NMR (125 MHz, $CD_3OD$) data are shown in Supplementary Table 1. Compound **10** (0.2 mg): HR-ESI-MS: *m/z* = 471.2602 $[M-H]^-$ (471.2600 calcd for $C_{24}H_{39}O_9$)**;** $^1H$ NMR (700 MHz, $CD_3OD$): $\delta_H$ 5.78 (1H, br s, H-2), 5.48 (1H, m, H-11), 5.47 (1H, m, H-10), 4.08 (2H, t, *J* = 6.6 Hz, H-7'), 4.02 (1H, m, H-5), 4.00 (1H, d, *J* = 5.6 Hz, H-7), 3.76 (1H, t, *J* = 6.2 Hz, H-6), 3.68 (1H, d, *J* = 11.6 Hz, H-16), 3.62 (1H, dq, *J* = 6.3, 4.9 Hz, H-13), 3.58 (1H, d, *J* = 11.6 Hz, H-16), 2.48 (1H, dd, *J* = 13.8, 4.0 Hz, H-4), 2.32 (1H, dd, *J* = 14.1, 8.4 Hz, H-4), 2.24 (2H, m, H-9), 2.22 (2H, m, H-2'), 2.21 (3H, s, H-15), 2.16 (1H, m, H-12), 1.66 (2H, m, H-6'), 1.62 (2H, m, H-3'), 1.35-1.45 (4H, m, H-4' and H-5'), 1.11 (3H, d, *J* = 6.3 Hz, H-14), 1.01 (3H, d, *J* = 6.9 Hz, H-17); $^{13}C$ NMR (125 MHz, $CD_3OD$): $\delta_C$ 181.1 (C, C-1'), 168.3 (C, C-1), 158.3 (C, C-3), 137.5 (CH, C-11), 126.6 (CH, C-10), 118.6 (CH, C-2), 86.9 (C, C-8), 81.5 (CH, C-5), 76.3 (CH, C-6), 75.8 (CH, C-7), 72.1 (CH, C-13), 65.1 ($CH_2$, C-16), 64.9 ($CH_2$, C-7'), 46.1 ($CH_2$, C-4), 45.4 (CH, C-12), 40.6 ($CH_2$, C-9), 37.3 ($CH_2$, C-2'), 30.3 ($CH_2$, C-4' or C-5'), 29.8 ($CH_2$, C-6'), 27.0 ($CH_2$, C-5' or C-4'), 27.0 ($CH_2$, C-3'), 20.4 ($CH_3$, C-14), 19.5 ($CH_3$, C-15), 16.6 ($CH_3$, C-17). Compound **16** (0.2 mg): HR-ESI-MS: *m/z* = 471.2591 $[M-H]^-$ (471.2600 calcd for $C_{24}H_{39}O_9$)**;** $^1H$ NMR (700 MHz, $CD_3OD$): $\delta_H$ 5.75 (1H, br s, H-2), 5.57 (1H, m, H-10), 5.48 (1H, dd, *J* = 15.5, 7.8 Hz, H-11), 4.08 (2H, t, *J* = 6.6 Hz, H-7'), 3.72 (1H, d, *J* = 2.9 Hz, H-7), 3.64 (1H, m, H-5), 3.62 (1H, m, H-13), 3.44 (1H, d, *J* = 10.9 Hz, H-16), 3.35 (1H, m, H-6), 3.32 (1H, m, H-16), 2.66 (1H, d, *J* = 14.4 Hz, H-4), 2.33 (2H, m, H-9), 2.26 (2H, t, *J* = 6.6 Hz, H-2'), 2.20 (1H, m, H-4), 2.18 (1H, m, H-12), 2.18 (3H, s, H-15), 1.66 (2H, m, H-6'), 1.62 (2H, m, H-3'), 1.35-1.45 (4H, m, H-4' and H-5'), 1.11 (3H, d, *J* = 6.4 Hz, H-14), 1.01 (3H, d, *J* = 6.9 Hz, H-17); $^{13}C$ NMR (observed via HSQC and HMBC): $\delta_C$ 168.4 (C, C-1), 159.0 (C, C-3), 137.4 (CH, C-11), 126.9 (CH, C-10), 118.3 (CH, C-2), 75.3 (CH, C-5), 74.1 (CH, C-7), 72.8 (C, C-8), 72.1 (CH, C-13), 70.9 (CH, C-6), 69.7 ($CH_2$, C-16), 64.7 ($CH_2$, C-7'), 45.7 (CH, C-12), 44.0 ($CH_2$, C-4), 39.8 ($CH_2$, C-9), 36.1 ($CH_2$, C-2'), 30.3 ($CH_2$, C-4' or C-5'), 29.8 ($CH_2$, C-6'), 27.0 ($CH_2$, C-5' or C-4'), 26.5 ($CH_2$, C-3'), 20.2 ($CH_3$, C-14), 19.4 ($CH_3$, C-15), 16.5 ($CH_3$, C-17).

**Gene Inactivation and Complementation of *mupZ.*** Targeted gene inactivation was achieved by the two-step allelic exchange method[28], as illustrated in Supplementary Figure 7. Two fragments sized 327 bp and 623 bp respectively upstream and downstream of *mupZ* were amplified from *P. fluorescens* NCIMB 10586 genomic DNA by PCR with the following primers (lower case letters indicate appropriate sequences homologous to target vector): *mupZ*-Up-For: 5'-acgaattcgagctcg-ATCCTAGCGTCCTCATCGTCT-3'; *mupZ*-Up-Rev: 5'-GTAGATACATCGGGTCCATTGAAGCTGAAACCGTAG-3'; *mupZ*-Down-For: 5'-TTTCAGCTTCAATGGACCCGATGTATCTACAAGAGGC-3'; *mupZ*-Down-Rev: 5'-ggccagtgccaagctCTACCAGGATCCAGACATTCAG-3'. The two PCR fragments were spliced together by overlap extension PCR using the upstream forward primer and the downstream reverse primer. The resultant 950 bp mutant allele was cloned into KpnI/HindIII-digested pEX18Tc by In-Fusion HD Cloning

Kit, yielding the allelic exchange suicide vector pEX18Tc-mupZKO, which was then transformed into *E. coli* S17.1 (λ *pir*$^{+}$). 1.5 ml of the overnight *E. coli* donor strain containing plasmid pEX18Tc-mupZKO and 0.5 ml of the *mmpEΔOR* mutant of *P. fluorescens* NCIMB 10586 recipient strain were centrifuged at 10,000*g* for 5 min at room temperature. Each cell pellet was collected and resuspended in 50 µl of LB and then combined. This cell mixture was transferred onto the middle of a prewarmed LB agar plate. The 'puddle' was allowed to dry on the agar surface and incubated overnight at 30 °C. The mating puddle was then scraped off and resuspended in 1.0 ml of sterile saline solution. 10-, 100-, and 500-µl aliquots were spread on LB agar plates supplemented with tetracycline (15 µg/ml) and carbenicillin (50 µg/ml) and incubated for 72 hours at 30 °C. Single colonies were picked from above LB plates and streaked onto no-salt LB + 15% (wt/vol) sucrose agar, followed by incubation for 48 hours at 30 °C. Colonies were screened for the targeted knockout mutant by PCR using the upstream forward and the downstream reverse primer pair (Supplementary Figure 7). For complementation, the *mupZ* gene was cloned into pEX18Tc using primers as follows: *mupZ*-Comp-For: 5'-acgaattcgagctcgATGAATCGCACCTGCATGGC-3'; *mupZ*-Comp-Rev: 5'-ggccagtgccaagctTCACGCCACCACAGACGG-3'. Plasmid pEX18Tc-mupZ was introduced into the *mmpEΔOR*/*ΔmupZ* mutant of *P. fluorescens* NCIMB 10586 by the same conjugation method as described above. Fermentation, isolation and purification of metabolites followed the same general procedure described previously[10]. All structures were confirmed by NMR spectroscopy.

**Crystallisation, data collection and structure determination of MupZ.** Crystals of MupZ were initially obtained using the sitting drop vapor diffusion method at 20 °C from the Morpheus II screen (Molecular Dimension Limited). Diffraction quality crystals were grown by microseeding using the hanging drop vapour diffusion method, mixing 1.0 µL of protein solution (5 mg/ml) with 1.0 µL of reservoir solution comprising 100 mM amino acids II, 0.1 M MOPSO, Bis-Tris pH 6.5 and 25% w/v PEG 4000, 40% w/v 1,2,6-hexanetriol. X-ray diffraction data were collected from a single crystal to a 1.45 Å resolution at beamline I03 of Diamond Light Source, UK. Data were indexed and integrated with DIALS[39] using xia2[40]. MupZ crystals belong to space group C121 with unit cell parameters *a*=42.33Å, *b*=80.09Å and *c*=78.82Å. The asymmetric unit contains 1 MupZ dimer of identical 14 kDa subunits with a solvent content of 39%. Initial phases were determined by molecular replacement with the program MOLREP[41] in CCP4i2[42,43]. A single subunit of the conserved protein MM_1583 from *Methanosarcina mazei* Go1 (PDB ID: 4DPO) was used as a search model. Prior to MR the model was truncated to remove any unstructured loops. Once a solution was found, about 80% of the model was automatically built using BUCCANEER[44] (CCP4i2). The resulting electron density allowed iterative rounds of model building and restrained refinement with in Coot[45] and REFMAC5[46] (CCP4i2). Data collection, phasing and refinement statistics are provided in Supplementary Table 2. In the final model of MupZ all the native sequence of the protein (2-122 residues) is clearly visible, with short section (VLFQGP) of the vector derived tag also present in subunit A. The refined coordinates have been deposited in the PDB with accession code 6FXD.

**Molecular docking.** Molecular docking of the expected epoxide substrate of MupZ was performed using the obtained crystal structure (without crystallographically determined water molecules). The substrate structure, truncated by -$(CH_2)_3$-$CO_2H$ for docking feasibility, was initially built with GaussView and optimized at the B97D/6-31+G(d,p) level in Gaussian09[47]; the C5-C6, C6-C7 and C7-C8 bonds were subsequently rotated to obtain a catalytically relevant conformation. AutoDockTools 1.5.4 was used to merge the non-polar hydrogens and define rotable bonds. During docking, all formally single rotable bonds were allowed to change, apart from the aforementioned C5-C6, C6-C7 and C7-C8 bonds. The side-chains of Glu54 and Asn73 were also allowed to rotate. Subsequently, docking was performed with AutoDock Vina 1.1.2[48], using a grid of 14×22×18 Å centered on the active site cavity. The exhaustiveness was set to 25. Docking poses where the truncated part was placed such that the -$(CH_2)_3$-$CO_2H$ could not feasibly fit in to the enzyme active site cleft were discarded. The top pose indicated a catalytically feasible orientation, and the -$(CH_2)_3$-$CO_2H$ was build back onto the substrate manually, in line with the protein structure (Fig. 5b). Based on the docked pose of the (truncated) substrate obtained, the system was set up for molecular mechanics simulation using the PREP protocol

of Enlighten (https://github.com/marcvanderkamp/enlighten/). This uses the AmberTools[49] program reduce to add hydrogens, and PropKa 3.1[50,51] to assign protonation states and His tautomers. All residues in their standard protonation states (in line with predicted pKa values by PropKa 3.1). All crystallographically determined water molecules were kept apart from water molecules in the active site cleft overlapping with the docked pose. Subsequently, the AmberTools program tleap was used to add additional waters to form a sphere of 20 Å around C5 of the substrate. Subsequently, the Enlighten STRUCT protocol was used for structural optimization (comprising of brief simulated annealing followed by energy minimization). The AMBER ff14SB force field for the protein[52], together with the TIP3P water model and General Amber Force Field parameters (GAFF)[53] for the substrate (assigned with Antechamber). The results from the structural optimization indicate a catalytically feasible substrate orientation, with Glu54 in line to abstract a proton from the hydroxyl group attached to C7, and Tyr41 interacting with the epoxide oxygen; Tyr41 could then donate a proton after (or concerted with) ring closure.

## Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request. X-ray crystallographic data are available in the EMBL-EBI PDB under the accession number 6FXD.

## References


1. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **79,** 629−661 (2016).

2. Carlson, J.C. *et al.* Tirandamycin biosynthesis is mediated by co-dependent oxidative enzymes. *Nat. Chem.* **3**, 628-633 (2011).

3. Bridwell-Rabb, J., Kang. G., Zhong, A., Liu, H.W. & Drennan, C.L. An HD domain phosphohydrolase active site tailored for oxetanocin-A biosynthesis. *Proc. Natl Acad. Sci. USA* **113,** 13750-13755 (2016).

4. Fuller, A. T. *et al.* Pseudomonic acid: an antibiotic produced by *Pseudomonas fluorescens*. *Nature* **234,** 416-417 (1971).

5. Thomas, C. M., Hothersall, J., Willis, C. L. & Simpson, T. J. Resistance to and synthesis of the antibiotic mupirocin. *Nat. Rev. Microbiol*. **8,** 281−289 (2010).

6. Martin, F. M. & Simpson, T. J. Biosynthetic studies on pseudomonic acid (mupirocin), a novel antibiotic metabolite of *Pseudomonas fluorescens*. *J. Chem. Soc., Perkin Trans. 1* 207-209 (1989).

7. El-Sayed, A. K. *et al.* Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. *Chem. Biol.* **10,** 419-430 (2003).

8. Hothersall, J. *et al.* Mutational analysis reveals that all tailoring region genes are required for production of polyketide antibiotic mupirocin by *Pseudomonas fluorescens*: pseudomonic acid B biosynthesis precedes pseudomonic acid A. *J. Biol. Chem.* **282,** 15451-15461 (2007).

9. Gao, S-S. *et al.* Biosynthesis of mupirocin by *Pseudomonas fluorescens* NCIMB 10586 involves parallel pathways. *J. Am. Chem. Soc.* **136,** 5501-5507 (2014).

10. Gao, S-S. *et al.* Selected mutations reveal new intermediates in the biosynthesis of mupirocin and the thiomarinol antibiotics. *Angew. Chem. Int. Ed.* **56,** 3930-3934 (2017).

11. Silvian, L. F., Wang, J. & Steitz, T. A. Insights into editing from an Ile-tRNA synthetase structure with $tRNA^{Ile}$ and mupirocin. *Science* **285,** 1074-1077 (1999).

12. Hemmerling, F. & Hahn, F. Biosynthesis of oxygen and nitrogen-containing heterocycles in polyketides. *Beilstein J. Org. Chem.* **12**, 1512-1550 (2016).

13. Taylor, R. D., MacCoss, M. & Lawson, A. D. G. Rings in drugs. *J. Med. Chem.* **57**, 5845-5859 (2014).

14. Luhavaya, H. *et al.* Enzymology of pyran ring A formation in salinomycin biosynthesis. *Angew. Chem. Int. Ed.* **54,** 13622-13625 (2015).

15. Cooper, S. M. *et al*. Mupirocin W, a novel pseudomonic acid produced by targeted mutation of the mupirocin biosynthetic gene cluster. *Chem. Commun.* **9**, 1179-1181 (2005).

16. Kauppi, B. *et al*. Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2-dioxygenase. *Structure* **6,** 571–586 (1998).

17. Sydor, P. K. *et al.* Regio- and stereodivergent antibiotic oxidative carbocyclizations catalysed by Rieske oxygenase-like enzymes. *Nat. Chem*. **3,** 388-392 (2011).

18. Suen, W. C. & Gibson, D. T. Isolation and preliminary characterization of the subunits of the terminal component of naphthalene dioxygenase from *Pseudomonas putida* NCIB 9816-4. *J Bacteriol.* **175,** 5877-5881 (1993).

19. Ferraro, D. J., Gakhar, L. & Ramaswamy, S. Rieske business: structure-function of Rieske non-heme oxygenases. *Biochem. Biophys. Res. Commun.* **338**, 175-190 (2005).

20. Li, B. *et al.* Whole-cell biotransformation systems for reduction of prochiral carbonyl compounds to chiral alcohol in *Escherichia coli*. *Sci Rep.* **4,** 6750 (2014).

21. Gally, C., Nestl, B. M. & Hauer, B. Engineering Rieske non-heme iron oxygenases for the asymmetric dihydroxylation of alkenes. *Angew. Chem. Int. Ed.* **54,** 12952-12956 (2015).

22. Jouanneau, Y., Meyer, C. & Duraffourg, N. Dihydroxylation of four- and five-ring aromatic hydrocarbons by the naphthalene dioxygenase from *Sphingomonas* CHY-1. *Appl. Microbiol. Biotechnol*. **100,** 1253-1263 (2016).

23. Kan, S. B. J., Huang, X., Gumulya, Y., Chen, K. & Arnold, F. H. Genetically programmed chiral organoborane synthesis. *Nature* **552,** 132-136 (2017).

24. Baldwin, J. E. Rules for ring closure. *J. Chem. Soc. Chem. Comm.* 734–736 (1976).

25. Gilmore, K. & Alabugin, I. V. Cyclizations of alkynes: revisiting Baldwin's rules for ring closure. *Chem. Rev.* **111,** 6513–6556 (2011).

26. Gilmore, K., Mohamed, R. K. & Alabugin, I. V. The Baldwin rules: revised and extended. *WIREs Comput. Mol. Sci.* **6,** 487-514 (2016).

27. Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K. & Jones, D. T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, W349-W357 (2013).

28. Hmelo, L. R. *et al.* Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nat. Protoc.* **10,** 1820-1841 (2015).

29. Holm, L. & Laakso, L. M. Dali server update. *Nucleic Acids Res.* **44,** W351-W355 (2016).

30. Hotta, K. *et al.* Enzymatic catalysis of anti-Baldwin ring closure in polyether biosynthesis. *Nature* **483,** 355-358 (2012).

31. Minami, A. *et al.* Allosteric regulation of epoxide opening cascades by a pair of epoxide hydrolases in monensin biosynthesis. *ACS Chem. Biol.* **9,** 562-569 (2014).

32. Capyk, J. K., D'Angelo, I., Strynadka, N. C. & Eltis, L.D. Characterization of 3-ketosteroid 9α-hydroxylase, a Rieske oxygenase in the cholesterol degradation pathway of *Mycobacterium tuberculosis*. *J. Biol. Chem.* **284,** 9937–9946 (2009).

33. Yoshiyama-Yanagawa, T. *et al.* The conserved Rieske oxygenase DAF-36/Neverland is a novel cholesterol-metabolizing enzyme. *J. Biol. Chem.* **286,** 25756–25762 (2011).

34. Li. J., van Belkum, M. J. & Vederas, J. C. Functional characterization of recombinant hyoscyamine 6β-hydroxylase from *Atropa belladonna*. *Bioorg. Med. Chem*. **20,** 4356-4363 (2012).

35. Mao, X-M. *et al.* Efficient biosynthesis of fungal polyketides containing the dioxabicyclo-octane ring system. *J. Am. Chem. Soc.* **137,** 11904-11907 (2015).

36. Minami, A. *et al.* Sequential enzymatic epoxidation involved in polyether lasalocid biosynthesis. *J. Am. Chem. Soc.* **134,** 7246-7249 (2012).

37. Shichijo, Y. *et al.* Epoxide hydrolase Lsd19 for polyether formation in the biosynthesis of lasalocid A: direct experimental evidence on polyene-polyepoxide hypothesis in polyether biosynthesis. *J. Am. Chem. Soc.* **130,** 12230-12231 (2008).

38. Barry, S. M. & Challis, G. L. Mechanism and catalytic diversity of Rieske non-heme iron-dependent oxygenases. *ACS Catal.* **3,** 2362-2370 (2013).

39. Parkhurst, J.M. *et al.* DIALS. *J. Appl Crystallogr*. **49,** 1912-1921 (2016).

40. Winter, G. *xia2*: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr*. **43,** 186-190 (2010).

41. Vagin, A. & Teplyakov, A. Molecular replacement with *MOLREP*. *Acta Crystallogr. D Biol. Crystallogr*. **66,** 22-25 (2010).

42. Potterton, L. *et al.* CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. D Struct. Biol.* **74,** 68-84 (2018).

43. Winn, M.D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr*. **67,** 235-242 (2011).

44. Cowtan, K. The *Buccaneer* software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Struct. Biol.* **62,** 1002-1011 (2006).

45. Emsley, P., Lohkamp, B., Scottc, W. G. & Cowtand, K. Features and development of *Coot*. *Acta Crystallogr. D Struct. Biol.* **66,** 486-501 (2010).

46. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst.* **D53,** 240-255 (1997).

47. Frisch, M. J. *et al.* Gaussian, Inc.: Wallingford, CT, USA (2009).

48. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* **31,** 455–461 (2010).

49. Case, D.A. *et al.* AMBER 2016, University of California, San Francisco (2016).

50. Sondergaard, C. R., Olsson, M. H. M., Rostkowski, M. & Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p*K*a values. *J. Chem. Theory Comput.* **7,** 2284–2295 (2011).

51. Olsson, M. H. M., Sondergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent treatment of internal and surface residues in empirical p*K*a predictions. *J. Chem. Theory Comput.* **7,** 525–537 (2011).

52. Maier, J. A. *et al.* ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11,** 3696–3713 (2015).

53. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25,** 1157–1174 (2004).

## Acknowledgements

We are grateful to BBSRC and EPSRC for funding through the Bristol Centre for Synthetic Biology (BB/L01386X/1), BBSRC for funding through BB/M012107/1 and for a David Phillips Fellowship (BB/M026280/1) to M.W.v.d.K. We thank Mara Malaysia for a scholarship to N. A. B. and Prof. Herbert P. Schweizer for kindly providing plasmids for gene inactivation and complementation.

## Contributions

L.W. and C.L.W. designed the experiments, analysed the data and together with M.P.C. and T.J.S. drafted the manuscript. L.W. conducted heterologous expression, protein purification, biotransformations, isolation and characterisation of metabolites and site-directed mutagenesis, A.P. and C.W. assisted with protein crystallisation and structure determination, M.R.C. assisted with cloning and protein purification, N.A.B. conducted substrate synthesis and model cyclisation, M.W.v.d.K. performed the molecular modelling, P.R.R. and M.P.C. led the protein structural studies and contributed to writing the manuscript.

## Competing Interests

The authors declare no competing interests.

## Graphical abstract